Arsenii Ashukha

Home page / Google Scholar / GitHub
e-mail: ars.ashuha@gmail.com tel: +79651782915

I'm a Research Scientist at Samsung AI Center Moscow. I (almost) received a PhD in Bayesian deep learning @bayesgroup so I can make big overcomplicated DNNs work :) My research interests are focused on crafting and training reliable and date&compute-efficient deep neural networks.

Prior to that, I was a part of Yandex Research team (a Russian search giant) and did ML engineering internships at Yandex and Rambler, where I worked on deep learning for audio and recommended systems. I received a master's degree in machine learning at MIPT and Yandex school of data analysis. I did a bachelor's degree at BMSTU with a major in applied math and computer science.

SELECTED PUBLICATIONS

Google Scholar: scholar.google.com/citations?user=IU-kuP8AAAAJ
Most representative works are highlighted

- ***Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning***, ICLR 2020 [arXiv]
  Arsenii Ashukha, Alexander Lyzhov, Dmitry Molchanov, Dmitry Vetrov
- ***Variational Dropout Sparsifies Deep Neural Networks***, ICML 2017 [arXiv]
  Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov
- ***Greedy Policy Search: A Simple Baseline for Learnable Test-Time Augmentation***, UAI 2020 [arXiv]
  Dmitry Molchanov, Alexander Lyzhov, Yuliya Molchanova, Arsenii Ashukha, Dmitry Vetrov
- ***The Deep Weight Prior***, ICLR 2019 [arXiv]
  Andrei Atanov, Arsenii Ashukha, Kirill Struminsky, Dmitry Vetrov, Max Welling
- ***Variance Networks: When Expectation Does Not Meet Your Expectations***, ICLR 2019 [arXiv]
  Kirill Neklyudov, Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov
- ***Uncertainty Estimation via Stochastic Batch Normalization***, *ICLR Workshop* 2018 [arXiv]
  Andrei Atanov, Arsenii Ashukha, Dmitry Molchanov, Kirill Neklyudov, Dmitry Vetrov
- ***Structured Bayesian Pruning via Log-Normal Multiplicative Noise***, *NeurIPS* 2017 [arXiv]
  Kirill Neklyudov, Dmitry Molchanov, Arsenii Ashukha, Dmitry Vetrov

CODE

- Check out very short and simple and fan to make implementations of ML algorithms:
  - Gradient Boosting
  - Real NVP normalizing flows
  - Quantile Regression DQN (Distributional RL)
- Also, check out more solid implementations:
  - Multi-GPU SimCLRv1 closely reproduced results on both CIFAR-10 and ImageNet
  - Ensembles (Deep ensembles, Snapshot ensembles, cSGLD, FGE, etc.)

EDUCATION

- *2017-2021* PhD in Computer Science, **Centre of Deep Learning National Research University HSE**. My PhD is focused on *applications and understanding of stochastic deep learning models*, Advisor: Dmitry Vetrov.
- *2015-2017* MSc in Computer Science**, Moscow Institute of Physics and Technology** jointly with **Yandex School of Data Analysis**, where I worked with Dmitry Vetrov on learning sparse DNNs.
- 2011-2015 BSc in Computer Science **Bauman Moscow State Technical University** with a double major in applied math and CS. I worked on linguistically motivated topic models with Natalia Loukachevitch.

PROFESSIONAL EXPERIENCE

- *Research Scientist*, **Samsung AI Center** (2018 - Now):
  Showed that modern uncertainty estimation techniques perform worse than the simplest 5 year old baseline, and widely used metrics for asses uncertainty estimation cannot provide a fair ranking of methods (results published at ICLR'20).

  The other topics I worked/working on include learnable data augmentation, learning prior distribution over DNNs' kernels with generative models, and self/semi-supervised learning. Since 2021 I'm helping to manage the lab.

- *Research Scientist*, **Yandex.Research & Centre of DL and Bayesian Methods NRU HSE** (*2016 - 2018*):
  Created Sparse Variational Dropout a method for sparsification of DNNs that, for the first time, allowed to achieve over 250x compression ratio (results published at ICML'17). The modified version of SparseVD with neuron-level sparsity allowed to accelerate inference of a CNN by 2-5 times and was applied to the feature extraction process for image retrieval (results published at NeurIPS).

- *Machine Learning Engineer Intern*, **Yandex.Music** (summer of 2016):
  I worked on feature extraction techniques for music data with convolutional neural networks. I also developed an evaluation of learned representations. The representations were used in the content-based recommendation system for yandex music.

- *Machine Learning Engineer Intern*, **Rambler&Co** (May - Oct 2015):
  Worked on demographic classification and recommendation systems. My responsibility included improving the quality and performance of classifiers, automatic feature extraction algorithms, and recommendation algorithms. Stack of technologies: Hadoop, Hive, Spark, XGboost, VW, gensim.

REVIEWING

- Conferences:
  - International Conference on Machine Learning, ICML (2019, 2020 top-33% highest-scored reviewers)
  - Neural Information Processing Systems, NeurIPS 2019 (top-50% highest-scored reviewers)
  - International Conference on Learning Representations, ICLR (2020, 2021)
- Workshops:
  - ICML Workshop on Invertible Neural Networks (2019, invertibleworkshop.github.io)

- Bayesian Deep Learning Workshop (since 2017, bayesiandeeplearning.org)

TEACHING

- Supervisor of reading clubs on machine learning at HSE and Yandex school of data analysis (since 2017)
- Talks and practical sessions at **Deep|Bayes** Summer School on Bayesian Deep Learning (since 2017)
- Lecturer, Moscow Institute of Physics and Technology: I was a lecturer and manager of the deep learning brunch of a facility-wide machine learning course ~60 students (ml-mipt.github.io). Also, I taught deep learning and practical sessions on cutting-edge ML algorithms on a facultative course "Data Mining in Action" ~ 200 students (https://bit.ly/3eRLGYp). The goal of this course is to make ML education available for everyone for free.

SUPERVISION:

- Alexander Lyzhov (moved to The Center on Long-Term Risk), Deep Neural Network Ensembles: Analysis and Approaches to Diversification (MSc, 2020)
- Andrei Atanov (PhD candidate at EPFL), Effective Learning of Deep Neural Networks Ensembles (BSc, 2018), Learning Deep Models with Small Data (MSc, 2020)
- Evgenii Nikishin (PhD candidate at Cornell), Stability Improvement and Knowledge Transfer in Deep Reinforcement Learning (MSc, 2019)

FRAMEWORKS & PROGRAMMING LANGUAGES:

- I'm fluent in Python and I use to code in C/C++, Go, language is not a problem after all.
- I'm also fluent with common Python libs such as NumPy, Matplotlib, scikit-learn, etc.
- My primary deep learning framework at the moment is PyTorch. Prior to that, I had a decent experience with Theano+Lagange and TensorFlow.
- I have experience with MapReduce, Hadoop, Hive, and Spark.